Repeatability and Reproducibility of NVG Gain Measurements Using the Hoffman ANV-126 Test Device

Denise L. Aleva
H. Lee Task
Crew Systems Interface Division
Air Force Research Laboratory
Wright-Patterson AFB, Ohio 45433-7022

Chuck Goodyear
Logicon Technical Services, Inc.
P. O. Box 317258
Dayton, Ohio 45437-7258

INTRODUCTION

Night vision goggles (NVGs) are being used extensively by our special operations forces for covert night operations. Often the individual operational units purchase the NVGs directly from the manufacturer. Upon delivery to the unit, the NVGs are tested to verify that they meet the gain specifications. The Hoffman Engineering ANV-126 portable test set is used for this purpose. However, the reproducibility of the NVG gain measurements obtained with the ANV-126 was unknown. Therefore, operators were uncertain as to whether to reject NVGs whose gain measurements were slightly below the criterion value. In addition, there was concern among NVG researchers and operators that the intensifier tubes in some NVGs might be greatly mismatched for gain, resulting in the luminance seen by the operator being significantly higher in one eye than the other.

Two specific objectives which are addressed herein were:
  • Objective 1: Determine the accuracy (repeatability and reproducibility) expected when measuring NVG gain using the Hoffman 126 test device.
  • Objective 2: Determine the distribution of Binocular Gain Ratios for fielded NVGs in both the linear gain radiance region and in the automatic brilliance control (ABC) radiance region of operation.
These objectives were addressed by an NVG gain data collection effort conducted at HQ AFSOC/LGMA NVG maintenance facility at Hurlburt Field, Florida. The testing took place during the period 18-20 November, 1996.

OBJECTIVE 1

If a test device is going to be used to make acquisition or acceptance decisions then it is necessary to determine the accuracy of the device to insure the validity of any resulting decision. The Air Force has acquired a number of ANV-126 NVG test devices but the reproducibility of the NVG gain measurements obtained with the device is unknown. Sources of variance expected are changes in gain of the NVGs themselves, the operators making the measurements, the test device itself and differences between test devices.

Method

In order to establish an estimate of the reproducibility of NVG gain measurements using the ANV-126 test device it was necessary to collect data on several NVG oculars using several test devices and operators. It was desirable to include as many test devices as possible to obtain a good estimate of the variance between devices. The test plan was designed in accordance with ASTM Publication E-691 (1992) which outlines procedures for testing repeatability and reproducibility.

Operators. The operators were five scientists from Armstrong Laboratory (Aircrew Training Research Division [AL/HEA], Mesa Arizona and Human Engineering Division [AL/HEC], Wright-Patterson Air Force Base, Ohio) and one representative from Hoffman Engineering. All had previous experience using the ANV-126 to measure NVG gain.

Test Sets. The test sets were nine Model ANV-126 Night Vision Device Test Sets for Ground Support Maintenance.

Report Documentation Page		Form Approved OMB No. 0704-0188
Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.		

1. REPORT DATE **1998**		2. REPORT TYPE		3. DATES COVERED **-**
4. TITLE AND SUBTITLE **Repeatability and Reproducibility of NVG Gain Measurements Using the Hoffman ANV-126 Test Device**				5a. CONTRACT NUMBER
				5b. GRANT NUMBER
				5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)				5d. PROJECT NUMBER
				5e. TASK NUMBER
				5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **AFRL,Crew Systems Interface Div,Wright Patterson AFB,OH,45433-7022**				8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)				10. SPONSOR/MONITOR'S ACRONYM(S)
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)
12. DISTRIBUTION/AVAILABILITY STATEMENT **Approved for public release; distribution unlimited**				
13. SUPPLEMENTARY NOTES				
14. ABSTRACT **see report**				
15. SUBJECT TERMS				

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**		**6**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

The ANV-126 is produced by Hoffman Engineering. It is a self-contained portable test set designed for field operational checks and depot-level NVG maintenance. Four of the test sets were resident at the Hurlburt facility, two were brought from Eglin Air Force Base, one from Hoffman Engineering, and one each from AL/HRA and AL/CFH. All test sets had been calibrated within one year of the test (range from 1 day to 8.5 months).

NVGs. The NVGs utilized in this testing were all F4949 D model. These NVGs were currently in use at Hurlburt Field by the SOF and were in the repair facility for periodic maintenance.

Procedure. Ten NVGs (20 oculars) were tested using 9 test sets and 6 operators (measurement personnel). NVG gain was measured using the standard procedures in the ANV-126 users' manual (1996) which include exposing the ocular to the maximum ANV-126 test level radiance input for 60 seconds prior to measurement. All gain measurements were made at the 0.1×10^{-3} fL radiance level.

Each operator tested 10 NVGs on each of 9 test sets. Six measurements were made of each NVG (3 for each ocular). Operators did not reset input level between measurements (i.e., the procedure was to exercise goggle, measure Left ocular, measure Right ocular, measure Left ocular, measure Right ocular, measure Left ocular, measure Right ocular, go to next NVG). It took an operator 30 - 45 minutes to measure all 10 NVGs with one test set. Operators by test set order was counterbalanced in case of learning or fatigue effects.

The day to day repeatability issue was addressed on Day 2. In this case we were concerned about the variability of the individual test sets over time. The procedure on Day 2 was the same as on Day 1, with the exception that 3 operators (Operators A, B, and F) tested each of 5 NVGs (all odd-numbered NVGs from Day 1) on each of 8 test sets (test set 9 was found to be defective). This, along with corresponding data from Day 1, allowed the assessment of repeatability of individual test sets over a 2-day period.

Results

Repeatability. Repeatability refers to the consistency of test results obtained by a single operator with a single test set. A repeatability limit was calculated for each of the 20 oculars. The repeatability limit tells us that for any pair of measurements by the same operator using the same test set and in a short time period, there is a probability of 0.95 that the second measurement will be within +/- the

repeatability limit of the first measurement. The repeatability limit is given by $Repeatability = 2.77\sqrt{\sigma_e^2}$ where σ_e^2 is the error variance. The average repeatability limit for the Day 1 measurements was 138 or 2.4 percent of the mean gain (overall mean gain = 5786. Repeatability was not found to increase or decrease significantly as a function of gain value (p = 0.67). Therefore, the repeatability of an individual gain measurement may be calculated by using the average repeatability value:

Repeatability limit = 138.

On Day 2, three operators measured 5 NVGs (10 oculars) using 8 test sets. Repeatability (same ocular, same operator, same test set) over Days 1 and 2 averaged 4.1 percent of the mean gain for all oculars. Again, repeatability was not found to increase or decrease significantly with gain value (p = 0.46); the calculated repeatability over the two-day period is given by:

Repeatability limit = 240.

Reproducibility. Reproducibility refers to the consistency of test results obtained by various operators using different test sets. A reproducibility limit was calculated in a manner similar to the procedure used to calculate the repeatability limit. However, the reproducibility limit is based on comparisons across all operators and test sets. The calculation of the reproducibility limit takes into account both operator and test set variances as well as the error variance for a single operator and test set. The reproducibility limit is represented by:

$$Reproducibility = 2.77\sqrt{\left(\sigma_s^2 + \sigma_o^2 + \sigma_{sXo}^2 + \sigma_e^2\right)}$$

where σ_s^2 is the test set variance, σ_o^2 is the operator variance, σ_{sXo}^2 is the variance of the test set by operator interaction and σ_e^2 is the error variance. Unlike the repeatability limit, which remained fairly constant across the range of gain values measured, the reproducibility limit was found to increase significantly with increased gain value (p = 0.0013); reproducibility limits for the 20 oculars ranged between 7.5% and 10.5% of the measured gain average. The reproducibility limit of an individual gain measurement is represented by:

Reproducibility limit
  = 186 + 5.6% of the gain measurement.

Discussion

While the repeatability and reproducibility limits do not give us an estimate of the accuracy of the gain measurement, they do allow us to calculate a range within which 95% of all differences between pairs of measurements can be expected to fall. For example, if our first measurement of an NVG ocular indicates a gain of 5800, then we can calculate the repeatability limit by: 5800 + 240 and 5800 - 240. This tells us that 95% of all second readings of that ocular by the same operator using the same test set within a 24 hour period can be expected to fall between 5560 and 6040. Likewise, we can calculate a reproducibility limit by: 186 + 0.056 * 5800 = 510.8. This tells us that 95% of all second readings by a different operator using a different test set can be expected to fall between 5289 and 6311.

Differences due to operators were much smaller than differences due to test sets. The gain values for test sets varied between an average of 5512 for test set 8 to an average of 5967 for test set 7. This is a range of 455. The average gain values for operators varied between 5765 for operator A and 5806 for operator C; this is a range of only 41. Thus, the range of values due to test sets is an order of magnitude greater than that due to operators. The highly trained operators appear to be responsible for relatively little of the variability in the gain measurements.

Difficulties encountered with test set 9 indicate that the test sets should be checked for proper functioning between routine calibrations. The ANV-126 users' manual should be consulted for the probe check procedure.

OBJECTIVE 2

There has been concern expressed by NVG researchers and operators that a large gain ratio (higher gain/lower gain) between the left and right NVG oculars may cause problems during NVG operations. Some evidence suggests that large gain ratios (greater that 1.5) may cause binocular rivalry and illusions of depth perception related to the Pulfrich Phenomenon (Pulfrich, 1922). There is currently no specification regarding gain ratio for NVGs. The purpose of this phase of the gain testing was to assess the extent of the problem. Gain ratios for a large population of NVGs were measured at Hurlburt Field, FL on 19 November 1996.

It is assumed that the binocular gain ratio (the gain of the NVG ocular with the highest gain to the gain of the other ocular) in the linear gain region is a reasonably accurate estimate of the binocular luminance ratio one might expect from an NVG. With this assumption, it is desirable to establish a maximum allowed binocular gain ratio to ensure that the luminance difference in the two channels of the NVGs will not be objectionable to the NVG user. In addition, it is desirable to measure the binocular luminance ratio of the NVGs when the tubes are in the ABC mode since this may be significantly different from the ratio obtained in the linear region. Again, the reason for this is to ensure binocular luminance compatibility. A separate study will determine what maximum allowed binocular gain ratio and ABC mode luminance ratio should be established. The purpose of this activity was to determine what levels of binocular luminance disparity exist in currently fielded NVGs. This activity had already been partially accomplished by Air Force Special Operations Command (AFSOC) which had collected gain data on 252 NVGs operating in the linear gain region. Table 1 is a summary of these data reduced to show binocular gain ratios: More than 97% of these NVGs had a gain ratio of 1.28 or less indicating the NVGs are reasonably well balanced in gain between right and left channels. These same data are shown in graphic form in Figure 1.

Table 1. Summary of AFSOC data for 252 NVGs showing percentage of NVGs achieving different levels of binocular gain ratios.

Gain Ratio	% of NVGs
1.04	19.4
1.08	37.7
1.12	55.6
1.16	77.8
1.20	89.3
1.24	92.9
1.28	97.2
1.32	97.6
1.36	98.8
1.40	99.2
1.44	99.6

Operators. The operators were three of the five scientists from Armstrong Laboratory (AL/HRA) who took part in the Objective 1 testing.

Test Sets. The test sets were 3 of the 9 Model ANV-126 Night Vision Device Test Sets for Ground Support Maintenance that were used for Objective 1. These test sets were identified as test sets 1, 2 and 3.

NVGs. Fifty-nine Model F-4949D NVGs were tested. These included the 10 NVGs used for the Objective 1 testing.

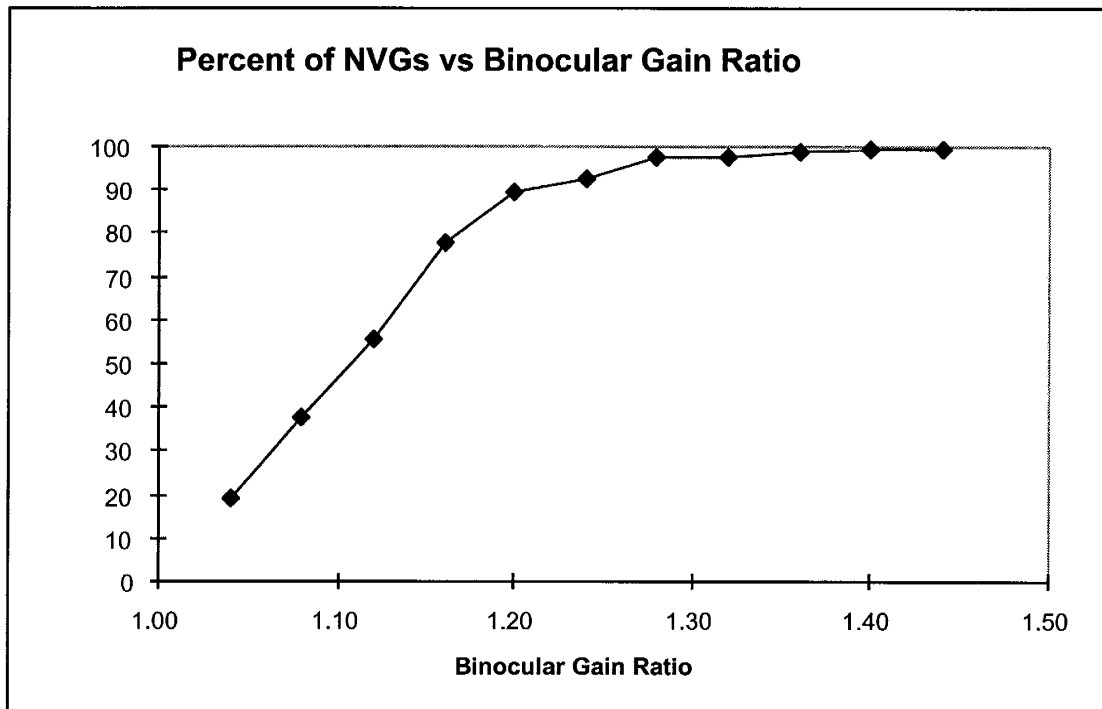**Percent of NVGs vs Binocular Gain Ratio**

Figure 1. AFSOC gain data showing percent of NVGs achieving various levels of binocular gain ratio. Only one NVG exceeded a ratio of 1.50 (it was 1.57).

Procedure. Each NVG was measured according to the procedures used to measure gain from the phase I testing. However, after gain was measured at an input illumination of 0.1 X 10^{-3} fL, the input level was increased to the maximum provided by the test set (approximately 1.4 X 10^{-3} fL). The illumination level placed the NVG in the automatic brilliance control (ABC) operating mode. Gain then was measured from each ocular under ABC conditions.

Results

The data were analyzed in a fashion similar to that shown in Table 1 and Figure 1. The results of the ratio measurements are shown in Table 2 and Figure 2 for the linear region and the ABC region. The results, in the linear region, revealed that only one NVG (s/n 3781 gain ratio = 1.72) had a gain ratio higher than 1.5. The average gain ratio was 1.11. In the ABC region, no NVGs had a gain ratio higher than 1.5, with the highest being 1.22 and the average being 1.06.

Repeatability and Reproducibility of Ratios. In a similar manner as described above, the repeatability and reproducibility of the NVG gain ratio between oculars was calculated. The repeatability limits (same operator, same test set) for gain ratio measurement can be calculated by:

Repeatability limit
    $= -0.037 + 6.4\%$ *of the gain ratio.*

Similarly, the reproducibility limits (different operator, different test set) can be calculated by:

Reproducibility limit
    $= -0.064 + 13.4\%$ *of the gain ratio.*

Discussion

These data permit a practical decision on establishing an acceptable level of binocular luminance disparity until data is available from other studies to determine if the interim binocular ratio should be changed (increased or decreased) based on visual performance and acceptance. Given the fact that many thousands of NVGs are currently fielded without significant user complaints regarding luminance disparity it is probable that the vast majority of these NVGs are in an acceptable binocular disparity range. The AFSOC data indicates that more than 97 percent of the NVGs they tested were at or below the binocular gain ratio of 1.3 which was the maximum level of disparity recommended in the Boeing Handbook for Equipment Design (Farrell and Booth, 1984). More than 93 percent of those tested in the current evaluation were at

436

or below 1.3 in the linear region and all were below 1.3 in the ABC region.

Table 2. Results of Gain Ratio Measurements of 59 NVGs in theLinear and ABC Regions.

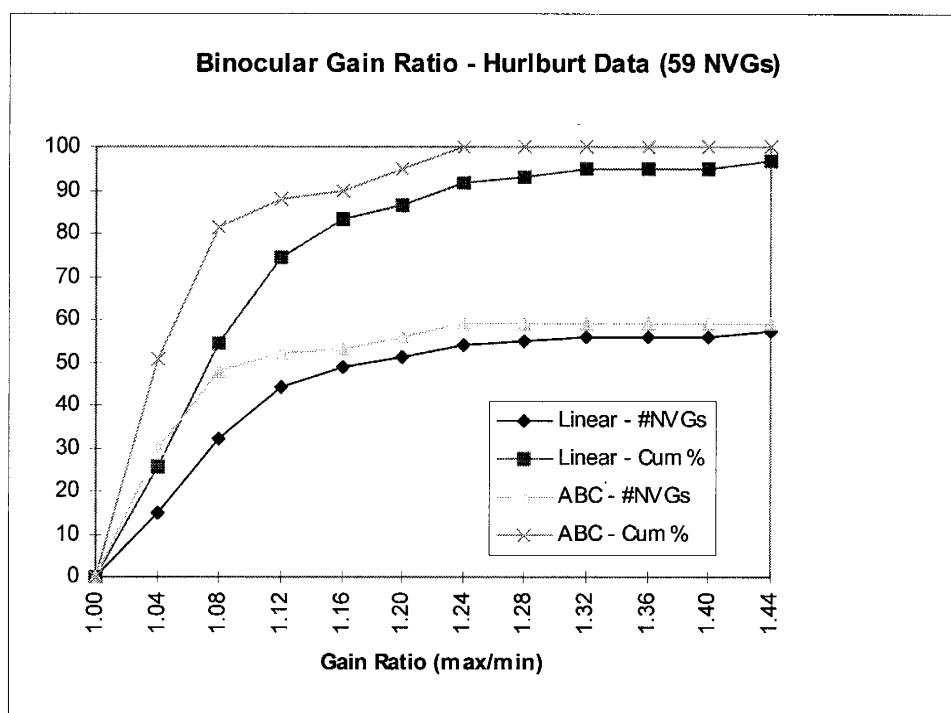| Gain Ratio	Linear - #NVGs	Linear - Cum %	ABC - #NVGs	ABC - Cum %
1.00	0	0	0	0
1.04	15	25.4	30	50.8
1.08	32	54.2	48	81.4
1.12	44	74.6	52	88.1
1.16	49	83.1	53	89.8
1.20	51	86.4	56	94.9
1.24	54	91.5	59	100
1.28	55	93.2	59	100
1.32	56	94.9	59	100
1.36	56	94.9	59	100
1.40	56	94.9	59	100
1.44	57	96.6	59	100

Figure 2. Results of gain ratio measurements of 59 NVGs in the linear and ABC regions.

As was discussed above for gain measurements, repeatability and reproducibility can be calculated for the measured gain ratio of a particular NVG. If, for example the measured gain ratio is 1.11, the repeatability can be calculated by: -0.037 + 0.064 * 1.11 = 0.034. Therefore, 95% of all differences between measurements of this NVG by the same operator using the same test set within a short period of time could be expected to fall between 1.076 and 1.144. Similarly, the reproducibility can be calculated by: -0.064 + .134 * 1.11 = 0.085. Therefore, 95% of all differences in the ratio measurement of this NVG by a different operator using a different test set could be expected to fall between 1.025 and 1.195.

CONCLUSIONS

The repeatability (same test set, same operator) and reproducibility (different test set, different operator) of gain measurements using the Hoffman 126 test device

437

have been quantified. These values should be useful to the operational units in setting limits for the acceptance/rejection of NVGs. However, this does not erase all uncertainty from the NVG acceptance/rejection question. Assume, for example, that the criterion for acceptance were gain equal to or greater than 5000. Table 3 gives the reproducibility limit and reproducibility range for several gain values around 5000. As we can see from the table, gain values between 4600 and 5400 result in uncertainty, as the reproducibility range spans both the unacceptable and acceptable regions. NVGs which test in this range may require further evaluation. However, we can feel confident that, for a criterion of 5000, NVGs which test below 4500 should be rejected and those which test above 5500 should be accepted.

Table 3. Reproducibility Limits and Ranges for a Selected Set of Gain Values

Gain Meas	Repro Limit	Repro Range
5500	494	5006 - 5994
5400	488	4912 - 5888
5000	466	4534 - 5466
4600	443	4157 - 5043
4500	498	4062 - 4938

It was determined that the gain ratio between oculars for

most of the NVGs tested did not exceed 1.5 in either the linear or the ABC region of operation. This finding was consistent with the fact that no NVGs were identified at Hurlburt Field as being regularly shunned by aircrew.

REFERENCES

ASTM Publication E 691 - 92. *Standard Practice for Conducting an Interlaboratory Study to Determine the Precision of a Test Method.* American Society for Testing and Materials. 1992.

Farrell, Richard J. and Booth, John M. *Design Handbook for Imagery Interrpretation Equipment.* Boeing Aerospace Company, Seattle, WA. February, 1984.

Model ANV-126 Night Vision Device Test Set for Ground Support Maintenance (Revision C). TM ANV-126, Hoffman Engineering Corporation, Stamford, CT. March 1, 1996.

Pulfrich, C. Die Stereoscopie in Dienste der isochromen und heterochromen Photometrie. *Naturwissen*, 1922, 10, 553-564.